Examining the Relationship between Cognition and Response Behavior in Health Retirement Survey

Yeabin Moon

May, 2020

Abstract

This paper examines the answering behavior of the old respondents. There is a growing interest in the cognitive decline of the aging population, but not enough is known about the consequences and implications of cognitive decline or its manageability from public policies outside the medical literature. The major challenge is insufficient measures of one's cognition even in popular surveys such as Consumer Expenditure Survey. The objective of the paper is to create the individual proxy from each respondent's answering behavior in the survey data. Note that taking a survey requires a series of cognitive tasks. Respondents need to comprehend the question, search their memory, make a decision, and format the response to a given scale. This paper proposes a standardized measure of answering behavior using open-ended financial questions in the Health and Retirement Study. I estimate the proxy of one's cognitive ability based on one's answering behavior and try to make it cross-sectionally comparable. The measures are in line with the respondents' cognitive score in the HRS.

1 Introduction

Cognitive decline is one of the most common conditions among the elderly. According to the Center for Disease Control and Prevention, every one in nine adults aged 45 and older is reporting experiences of cognitive decline. It is also one of the earliest noticeable symptoms of Alzheimer's disease and related dementias Jessen et al. (2014) and Facts and 2018 (2018). Cognitive decline can necessitate long-term care, creating medical costs for individuals in cognitive decline. There is wide research interest among economists in understanding the causes and the consequences of cognitive impairment. However, only a limited set of nationally representative surveys include a cognition measure, even fewer of those ever follow the same individuals over time. This lack of data availability hinders researchers' ability to examine questions of high policy importance.

In this paper, we aim to create a cognitive index based on the way an interviewee answers survey questions. In particular, we utilize how an individual rounds numbers when responding to open-ended financial questions and how frequently they skip questions to form the index. We hypothesize that with cognitive decline or dementia, a person could lose precision in numbers and provide fewer significant figures in their answers to open-ended financial questions, or they may be more likely to choose not to answer a question.

We test our hypothesis using the Health and Retirement Study (HRS), a nationally representative longitudinal survey. It includes several questions testing respondents' memory and cognition, and it also contains some open-ended financial questions. We construct our index based on the individuals' rounding patterns in responses to the latter, which include questions about home value, property tax, checking account balance, food expenditure, transportation, social security income, and drug/doctor/dentist expenditure. We also incorporate whether individuals answer "do not know", "refused to answer," or "valid skip" in the construction of this index. In our analysis, a few interesting patterns emerge. First, there is a clear pattern with age. An older individual is more likely to round (have fewer significant figures) and/or not to provide an answer (conditional on education, marital status, wealth, and interview method). The relationship between age and our proxy is statistically significant regardless of whether or not we include individual fixed effects. There is also a positive correlation between cognitive ability (measured by HRS) and our constructed index (conditional on education, marital status, wealth, interview method, and age). Again, the result is robust to the inclusion of individual fixed effects. The robustness of our results to the inclusion of individual fixed effects suggests that our finding does not solely rely on between-individual variation in answering/rounding behaviors.

To further validate whether the index we construct represent cognition beyond the HRS dataset, we utilize Panel Study of Income Dynamics (PSID). In PSID, a similar set of open-ended financial questions was included. PSID only survey cognition in 1972 of the household heads in their panel of over 50 years. Therefore in order to validate our method, we examine a well-known stylized fact that 1919 birth cohort which had in utero exposure to 1918 Influenza Pandemic had worse cognition (education outcomes), as found in Almond's seminal paper Almond (2006). We check whether we can observe the pattern found by Almond using the index we created in PSID, and we did find that.

2 Literature and Relevance

Rounding true values to even multiples of reported units is called heaping, and it is frequently observed in survey data. Statistics and econometrics literature try to estimate the degree of heaping in the data and develop techniques to handle resulting problems such as attenuation bias (Manski and Molinari (2010), Roberts and Brewer (2001) and Zinn and Würbach (2016)). Our focus is the answering behavior of the open-ended financial questions in survey data, and we try to infer the respondent's cognitive ability from the answering pattern. It can be easily applicable to any other numerical questions in other surveys, and the response in the financial questions less likely reflects the true distribution (Riddles et al. (2017) and Battistin, Miniaci, and Weber (2003)).

By carefully classifying the respondent's behavior, our study adds to growing literature on related issues. The related body of work, suggested by Krosnick (1991), empirically analyzes whether the change in cognition is correlated with how much the respondents complete the survey. One set of papers studies why the respondents choose opt-outs. Many surveys offer the respondents explicit optout options such as "don't know" or "refused to answer". Colsher and Wallace (1989) show that the old respondents more likely chooses opt-outs to particular type of questions. Knäuper et al. (1997) study how often the respondents aged 70 years old and older choose "don't know" and find that the respondents with low cognitive ability tend to choose it more on difficult questions. Another set of paper examine the relationship between the cognitive ability and the degree of detail about the numerical responses. Holbrook et al. (2014) focus on the numerical response, which is divisible by 5. Gideon, Helppie-McFall, and Hsu (2017) examine the rounding behavior in a broader manner.

They study intraclass variation of the respondent's rounding behavior. Andrews and Herzog (1986) examine the relationship between the variance of answers and the age of respondents in popular surveys. Previous literature suggests that answering behavior is correlated with the cognitive ability. Then, our question is whether we could create an numerical value based on the answering behavior as a proxy for the cognitive ability.

3 Data Description

We use the 2004 to 2016 waves of Health and Retirement Study, a nationally representative longitudinal survey of Americans aged 50 and older and their spouses or partners. It provides the respondents' health status, household portfolios and family structure. We use the 2004 survey as the starting wave because it is the first year in which the baby boomer generation is newly added. Most of questions are asked at the individual respondent level, but some questions are asked at the household level. For two-respondent households, household level questions are asked of one respondent who is labeled as the financial respondent, family respondent, or coverscreen respondent (the first respondent interviewed) on behalf of the entire household. In most cases, one respondent handles multiple roles. Because our interest lies on the open-ended financial questions, we mainly focus on the individual who responds the financial questions as a representative of their family members[1]. In case the financial respondent is changed from the past year, we treat the prior individual as missing in the current year. Also, we exclude the respondents who have any type of assisting helper.

As a result, pooling 15,729 individuals across the seven waves, we construct a sample of 55,571 individual-years. Table 1 presents the parsimonious summary statistics of the sample. The panel is not balanced – only about 15 percent of our household present every wave. However, a balanced panel is not required for our study. The median number of waves that each respondent participate is three. The number of female respondents are dominating, and the female proportion grows as the age of respondents increase. On average, the respondent graduate high school, and many respondents change the mode of interview year by year.

1. We only examine respondent-level files and household-level files. That is, we ignore sibling level, household member and child level, and helper level questions

3.1 Cognitive score

The HRS has a series of tests on one's cognitive ability. The set of questions in section *D cognition* are related to episodic memory, mental status, and comprehension of vocabulary. We define the cognitive score as the sum of three questions *immediate word recall*, *delayed word recall*, and *serial 7's test*, and use it as a reference index of the cognitive ability. The set of questions appear in every wave without any modification of the question statement, and all the respondents have an access to these questions regardless of their age. Furthermore, these questions are mostly relevant to verbal learning, reasoning, and attention abilities (Herzog and Wallace 1997).

For the immediate word recall, ten list of nouns are given to the respondents, and then they are asked to recall the list in order. A few minutes later, the respondents are asked to recall the same list of the words again. This is the delayed word recall. For the serial 7's test, the respondents are asked to subtract 7 from 100, and continue subtracting 7 from the last response. The respondents are asked to try this five times. The score range of the immediate word recall and delayed word recalls are both zero and ten, and the serial 7's test is between zero and five. Therefore, the cognitive score ranges from zero to twenty-five[2].

Figure 1 shows the age trend of the cognitive scores and its compositions. Although the magnitude is different among four panels, they show the similar trends. There is little variation until age 65, but the sharp declines are observed after that. There are no consensus when the cognitive decline starts in medical literature, but there is a general agreement that the cognitive decline would be accelerated for the old people (Karr et al. (2018)). Hence, we limit the respondent's age from 65 to 90 and benchmark our performance of the proxies against the cognitive score. In appendix[3], we present the trends of the cognitive score by gender and education level. Female respondents tend to have higher cognitive scores, and highly educated respondents tend to have higher cognitive scores.

2. Note that the response rate is somewhat different among the measures. We construct the cognitive score if the respondents answer all three tests.

3. See Figure 11.

3.2 Question selection

We restrict our analysis to a set of topics related to monetary questions [4]. From each section, we first select open-ended numerical questions. This rules out closed-ended questions and time related questions. Then, we select the questions based on the response rate. Ten questions meet the criteria and appear in every wave: the present value of home, real estate taxes, social security income, checking account, transportation value, food at home, food away from home, out-of-pocket for doctor visit, dental bills and drug costs. Figure 2 presents the average of numerical response by age for each question. One common characteristics among the ten questions are the strong patterns of heaping on round numbers (Figure 3).

Table 2 presents the summary statistics for the ten questions. The variance of the responses are huge, and the opt-out rates are quite high. We exclude the question of expenditure on eating out since more than 40 percent of response is a single digit [5]. Also, we exclude home and vehicle value questions. The numerical answers on these questions are based on the respondent's judgment and memory. The transactions occur far less frequently compared to other items, and hence it is very unlikely that the respondents provide the detailed numerical value. We examine these questions in the robustness-check subsection.

3.3 Demographic variables

Demographic variables include age, mode of interview[6], gender, years of schooling, and wealth. These are important along with the cognitive score to test our proxies are consistent with previous findings. We define total wealth as the sum of nonfinancial wealth, retirement wealth, and other financial wealth. Since the values are nominal, we first deflate the values using the price index in 2012 for Gross Domestic Product from the National Income and Product Accounts. The years of schooling does not change by time. It is possible that the respondents gain more schooling between the survey years. However, only four respondents in our sample gain more years of schooling. Race and where

4. Financial questions mostly belong to the following sections: health care cost (N), family structure (E), housing (H), asset and income (Q), capital gains (R), job (J, L), disability (M), insurance (N), and divorce (S).

5. There is no trailing zero for single digit response.

6. The mode of interview indicates whether a respondent is interviewed by either face-to-face or phone.

the residents live are probably important factors, but these are restricted information in the HRS.

4 Characterizing Response Behavior

In this section, we characterize the respondents' answering behavior. Although open-ended questions cannot be answered with a statistic response such as "yes" or "no", the HRS provides respondents the opt-out options "don't know" and "refused to answer". Hence, characterizing the behavior takes account the degree of details on numerical responses and whether the one chooses opt-outs. We start with measuring the degree of details for each question and consider how to add opt-out response subsequently.

4.1 Measuring the degree of details

We count the number of non-trailing zeros and the total number digits for each numerical response, and then create the index of the degree of details. We introduce two ways of creating the index which look similar but present subtle differences. We define m as the total number digits and n as the number of non-trailing zeros. The first index is following Gideon, Helppie-McFall, and Hsu (2017), in which the index is calculated by $\frac{n-1}{m-1}$[7]. The second index is calculated by $\frac{n}{m}$. 1 is the least upper bound for both indexes, and the larger value means the more detailed response.

There are two differences. Consider two responses \$1,000 and \$10,000,000. Following the above notation, n is equal to 1 for both responses, m is equal to 4 and 8 respectively. So, the firsts index calculates both responses as zero. Gideon et al. label this type of responses as *maximal rounding*. On the other hand, following the second index results in 0.25 and 0.125, and therefore the latter response is regarded as less detailed. That is, the scale of response matters for the second index especially when the respondents do maximal rounding. The second difference is the range of the index. Although both measures use the fractions less than or equal to one, $\frac{n}{m}$ does not include zero in the range since n never be zero for numerical responses.

7. The original method by Gideon et al. is $\frac{m-n}{m-1}$. We transform its direction using $1 - \frac{m-n}{m-1}$ to make two measures are comparable in direction.

There are two findings presented in both indexes. First, although the range of the fractions could be large, but the certain fractions are much more observed than others. Figure 4 presents the distributions of the first index among the selected questions. The significant mass is observed on zero, maximal rounding, in most questions. The similar patterns are observed in the second index[8]. Although the distribution looks more spread, only a few fractions have the observations. The second finding is that the large number bias would not be a concern. Some might argue that one would respond with more trailing zeros if the magnitude of the response is larger. Figure 5 presents the local average of numerical response by $\frac{n-1}{m-1}$ the first index. If there were large number bias, the negative relationship would be expected. However, Figure 5 shows no linear patterns. The second index presents the similar result.

4.2 Constructing Proxy

The HRS offers opt-outs in most questionnaires. The response rate of opt-outs in the HRS is quite high (Table 2)). Ideally, the proxy should reflect the varying degree of details on numerical responses. At the same time, it would be better if the proxy can capture whether the respondents choose opt-outs, because too many respondents often choose it. We build the proxies based on the two index we constructed above. In the rest of paper, we discuss seven proxies constructed in the following manner. We define proxy 1 as the average of the first index for the selected questions, and proxy 2 as the average of the second index for the same set of questions. That is, proxy 1 and proxy 2 only deal with the degree of details on numerical questions. The rest of proxies explicitly take account the opt-out option. To construct proxy 3, we calculate the index $\frac{n-1}{m-1}$ for each question. If the respondent answers DK, we assign 0 to the index. Proxy 3 takes the average of them. Proxy 4 is similar to proxy 3 but uses the index $\frac{n}{m}$. The interpretation between proxy 3 and proxy 4 has a subtle difference. Proxy 3 treats DK as the same as the maximal rounding. On the other hand, in proxy 4 DK means that the respondent answers with the less detailed degree compared to any numerical answer.

Note again that certain fractions are much more observed in Figure 4 in the first index. It raises the concern on proxy 3 and proxy 4 because the most variation in the proxies might come from

8. See appendix Figure 12.

whether one chooses the opt-out or not. To test this concern, proxy 5 and proxy 6 focus on whether the respondent do maximal rounding. To construct proxy 5, for each question, we construct the index in the following way. 0 is assigned to DK, 1 is assigned to maximal rounding, and 2 is assigned to any other numerical values. Proxy 5 takes the average of them. To construct proxy 6, we assign zero to either DK or maximal rounding, and one to any other numerical responses for each question. Proxy 6 takes the average of them. It is possible that the useful variation only comes from whether the respondents choose opt-outs or not. Proxy 7 tests this claim. For each question we assign zero to DK and one to any numerical response, and then take the average of them.

5 Proxy Evaluation

In this section, we evaluate the proxies based on the two aspects. Note that Figure 1 shows the average cognitive score declines as age increases. The minimum condition of the proxy is whether it can capture the aging pattern for the respondents ages 65 and older. Interestingly, HRS has a panel dimension so that we can further examine whether the individual over-time pattern of proxies match the corresponding cognitive score. Secondly, we assess the correlation between the proxies and demographic variables to check whether the results are consistent with the previous literature findings. Each proxy is based on the distinct counting method. From the regression exercise, we can also examine whether the results are sensitive to the counting methods. Since the range of cognitive score and proxies are different, we standardize the cognitive score and all proxies.

5.1 Local average of Proxies by age

Figure 6 presents the local average of proxies by age. The cognitive score in the first panel serves as a benchmark. Note that the proxy 7 only deals with whether the respondents choose opt-outs or not. It is quite simple but presents the consistent pattern with the cognitive score according to the last panel. Then, the question is whether incorporating rounding behavior provides a marginal value to predict the cognitive ability. Also, given that many surveys do not provide explicit opt-out options, how much proxy 1 and proxy 2, which takes account the rounding behavior only, provides the information on cognitive ability should be assessed. All proxies except proxy 1 and 2 seem to capture the age pattern of cognitive score closely. Proxy 1 and proxy 2 show the decreasing trend with high variance.

Since the HRS is a panel data, we can examine the proxy pattern within the same person by controlling for the individual fixed effects. Figure 7 presents the local average by age, and the first panel serves as a benchmark again. All the panels have the negative trends, and even the proxy 1 and 2 correspond to the pattern of the cognitive score. Note that in both figures proxy 3, 4, 5 and 6 show very similar patterns compared to the pattern of the cognitive score. It might imply that the resulting patterns are not sensitive on how to incorporate opt-outs. On average, we find that respondent not only rounds more but also opts out more as they get older. Taking account rounding behavior only would estimate the cognitive ability poorly for the cross-sectional data. However, in the panel data proxy 1 and proxy 2 would be a good proxy for the cognitive ability[9].

5.2 Empirical Analysis

The empirical framework models the extent to which answering behavior varies with the characteristics of the respondents. In the regression analyses, we try to find correlations and associations among variables that could be helpful in understanding the respondents' answering behavior. We test all proxies and examine the partial correlation between each proxy and demographic variables. Again, the HRS has the cognitive score so that we can examine whether our results match the previous findings in literature, and therefore we might have a partial justification for our proxies. Of course, we acknowledge that the estimation risks omitted variable bias, but we mitigate this concern in part by providing ample robustness checks. In this section, we test whether response heaping is more observed among the low-income respondents (Myers 1976). Using the years of schooling, we examine the claim that illiterate respondents present more response heaping (Budd and Guinnane 1991). Lastly, we test whether female rounds more as in Boyle and Gráda (1986). All the previous findings are based on the cross-individual analysis. We examine which proxies are consistent with the literature and further explore whether the results are changed in the individual fixed effect models.

9. It would raise the questions on whether the results depend on the question selection. We discuss it in the robustness-check section.

We estimate several variants of the following linear regression model for the respondent i at time t,

$$Proxy^j_{i,t} = \alpha^j_i + \beta^j_1 age_{i,t} + \beta^j_2 Interview\ mode_{i,t} + \beta^j_3 wealth_{i,t} + \beta^j_4 cognition_{i,t} + X_{i,t}\gamma + \epsilon^j_{i,t} \tag{1}$$

where $j = \{1, 2, 3, 4, 5, 6, 7\}$. i and t indicate the survey respondent i and survey year t. j denotes the classification of the proxy. age is the respondent's age, $Interview\ mode$ indicates the method of survey[10], $wealth$ is wealth quantile of total wealth, and $cognition$ is the cognitive score in the HRS. $X_{i,t}$ is a covariate vector including gender and years of schooling. So far, we examine relatively parsimonious specifications with very limited number of demographic variables. Gender and years of schooling do not change over time within the same individual, and hence in the individual fixed effect model, γ is not identified. $\beta^j_1, \beta^j_2, \beta^j_3$, and β^j_4 are parameters to be estimated. In most surveys, variables on cognition are not available, but we add it into the model because the partial correlation between the proxies and cognition in part determines their justification for usage.

The age effect is captured by a continuous age variable. Since the respondent's answering behavior is a function of age, controlling for age effect is important in order for other parameters to capture the net effect on the behavior. The mode of interview is included to account for difference in answering attitudes. The wealth quantile is included to examine whether the pattern of answering behavior is different based on economic conditions. We expect the estimate of β_1 to be negative and significantly different from zero, while the estimate of β_2 to be much smaller in magnitude or statistically insignificant. The estimates of β_3 and β_4 are expected to be positive.

Table 3 presents coefficient estimates for the equation (1) without the individual fixed effects. Each column shows the resulting coefficients using the proxy as a dependent variable. Note that the sign of the coefficients on age, cognitive score, and mode of interview are very similar among the proxies. On average, the respondents more likely round the number (or choose opt-outs) when the respondents are older controlling for cognition and other demographic variables. Similarly, controlling for age and other variables, the respondents on average less likely round the number (or less likely chooses opt-outs) when they have higher cognitive ability.

10. It is a binary variable. The value one indicates the phone interview, and zero indicates face-to-face interview.

Note that there is no conclusive result on the effect of gender, years of schooling, and wealth. The sign and magnitudes of coefficients on education, female and wealth differ even among proxy 4, 5, 6 and 7 in contrast to Figure 6 in which proxy 4, 5, 6 and 7 present similar patterns. Boyle and Gráda (1986) find that women present more response heaping than men, but the consistent results are only found in proxy 5 and proxy 7. Also, Budd and Guinnane (1991) show that response heaping is more frequently observed among illiterate respondents, and hence the sign of education might be a concern for all proxies except proxy 6 although the correlation between the years of schooling and illiteracy is somewhat vague. Lastly, Myers (1976) finds that the low-income respondents more likely show response heaping, but the results are incompatible among the first four proxies in our regression.

There are two takeaways from the cross individual analysis. First, the regression results are sensitive to the counting methods. Note that we use two similar methods measuring the rounding behavior for each question. Proxy 1 is based on $\frac{n-1}{m-1}$, and proxy 2 is based on $\frac{n}{m}$. They look similar but have different inference on maximal rounding. The column (1) and (2) present that the resulting coefficients have different magnitudes and signs. Secondly, the regression results are sensitive to how to incorporate the opt-outs. We implicitly make two separate assumptions to add opt-outs to the proxy. First assumption is that DK means more rounding than maximal rounding. This assumption is applied to proxy 4, and proxy 5. The second assumption is that answering DK and maximal rounding are considered the same. Proxy 3 and proxy 6 are based on this assumption. Especially, the only difference between proxy 5 and proxy 6 is how to handle the maximal rounding. Column (5) and (6) show the results on proxy 5 and 6 and present that the resulting coefficients have different magnitudes and signs. The regression results from the cross individual analysis might indicate that using a single proxy is not enough to estimate the cognition. However, as in Figure 6, each proxy could provide meaningful information on the cognitive ability, and hence combining proxies might be necessary to estimate one's cognition for the cross individual data. We test this by running the regression of the cognitive score on proxies in Table 4. The column (1) and column (2) only contains the single proxy. The column (3) includes multiple proxies and show the cognitive score is better explained compared to the single proxy model. In fact, the proxy selection is the result of Lasso adaptive model. We only test 7 proxies along with the demographic variables. We leave the complete model selection by Lasso for future works.

Table 5 presents the coefficient estimates for the equation (1) with the individual fixed effects. On average, controlling for the change in cognitive score and other demographic variables, the respondents are rounding more (or opt out more) as they get older. As in Figure 7, controlling individual fixed effects improves fits dramatically. Among proxy 3, 4, 5, 6 and 7, all the signs of the coefficients are the same, and the magnitudes are also similar. Even proxy 1 and proxy 2 present the similar estimates except wealth. Compared to the cross-individual analysis, in the panel data, either counting methods or how to incorporate opt-outs would not affect the estimation much. Also, given the variation in proxies are largely explained by the variation in cognitive ability, each proxy can be a good estimate of the respondent's cognition in panel data.

Note that gender do not change within the same person over time, and hence it is omitted in the fixed effect models. Note also that two groups are hardly comparable as the number of observations denotes. The size of female households grows more and more as the respondents get old. In appendix[11], we present the individual fixed effects model by gender. We do not find meaningfully different implications compared to Table 5.

5.3 Robustness check

Note that the cross individual analysis is sensitive to counting methods and how to incorporate opt-out responses, and hence in this subsection, we perform the robustness checks on the individual fixed effect models only.

We use 7 questions and take the average of the responses according to the construction method for each proxy. One might worry that one or two groups of responses influence greater than others so that they would bias the resulting averages. To check this conjecture, we leave one question away and take the average of responses from the remaining six questions. Following the construction rule for proxy 4, we construct seven different proxies. Table 6 presents the coefficient estimates. The first column label 'Pro tax' means that the property tax question is left out. Interpretation on the rest of columns follows the first column. Compared to column (4) in Table 5, we find no notable difference. We try the same robustness checks on all other proxies but find no meaningful difference as well.

11. See Table 12 and Table 13.

Another possible concern would be that we use too small number of questions. Now we add the home value question and the vehicle value question. We rule out them because they are too subjective questions. Unlike checking account, there are no convenient references, and therefore answering detailed number is hardly likely. Table 7 present the result, and the specifications are the same as Table 5. We find no significant difference.

One might raise the question whether we can find the similar findings when we extend the sample age. Now we include the respondents ages 50 and older and estimate the parameters using the same set of questions as Table 5. Table 8 presents the results. Note that the signs of age are inconsistent among the proxies. It is due to the fact that the cognitive ability does not change much until age 65 (Figure 1). Hence, it is possibly advisable to use the proxies for the respondents ages 65 and older.

One of the most challenging concern is to deal with the respondent's motivation. According to Krosnick (1991) and Gideon, Helppie-McFall, and Hsu (2017), rounding is more common for respondents who are low in motivation. Thus far, we have not controlled anything related to motivation, and hence our estimates might have an omitted variable bias. Motivation could be manifested in a various form. To mitigate the issue, we try to construct an index as a proxy for motivation. We note that some respondents choose DK dominantly in most questions (Figure 9). It is possible that the respondents overlook the questions and choose DK unconsciously because of low motivation. Here we use the ratio of opt-outs to the total number of responses as a proxy for motivation. That is, the less motivated respondents tend to choose more opt-outs.

We use the all the questions in three sections[12]. We calculate the proportion of opt-outs to all the responses and label it as 'demotivation'. Table 9 presents the individual fixed effects model using each proxy as the dependent variable. As expected, the sign of demotivation is negative and statistically significant. Controlling for the change in age, cognition and other controls, more de motivated respondents tend to round more (or more likely chooses opt-outs). Compared to Table 5, we find there is no meaningful difference in terms of the magnitude and significance, and hence we conclude that motivation might not be a concern.

12. The selected sections are 'Health Care Costs', 'Housing', and 'Assets, Debts, Income'. 7 questions selected for proxies belong to these sections.

6 Future Works: Application to the PSID

The goal of the project is to construct the index from the way people respond to survey questions even when the direct measure of cognition is not available. There are several reasons why we choose PSID as the application. First, it has a panel dimension. Second, it offers the explicit opt-out options DK. Lastly, PSID has the similar set of questions as in the HRS. We examine the latest five waves of data from 2009 to 2017.

The PSID was launched in 1968 by the Institute for Social Research at the University of Michigan. It was annually published until 1997, but since then it has been published once in two years. The PSID provides detailed information on assets and liabilities with diverse demographic backgrounds. The PSID respondents consist of three groups: the core sample, low-income-family sample, and immigrant sample. The core sample has the largest proportion of them, and they are sampled from the Survey Research Center (SRC), which are representative of the US population. The initial criteria for sample selection focuses household head sampled from the SRC. We select the respondents ages 65 and older.

The list of questions is property tax, checking account, social security income, doctor visit expenditure and drug expenditure. We construct 7 proxies and standardize them as in the HRS analysis. PSID has richer set of demographics compared to the HRS. Information on race and state-level geocode is available even in public data. The downside of data for our purpose is too few observations on old respondents (about 780 respondents in each wave). Figure 9 presents the local average of proxies by age. It is puzzling that there are opposing trends among the proxies. Figure 10 shows the local average controlling for the individual fixed effects. There is no upper sloping, even though the trends are not as clear as the HRS data in Figure 7.

Next, we examine the correlation between the constructed proxies and the respondents' characteristics. We regress each proxy on age, gender, education, marital status, and wealth quantile[13]. The cross individual analysis presents in Table 10. Like Figure 9, the resulting coefficients are not consistent among the proxies. Further, the sign on age is not intuitive. Table 11 presents the individual fixed effects model showing more consistent patterns among the proxies. Thus far, it is less clear whether the inconsistent results are due to the data characteristics or the methods. The methods can

13. Interview mode is available, but all the respondents in the sample use the phone survey.

be tailored to meet data characteristics. Again, the analysis is preliminary, and hence we have not concluded whether our method is applicable to PSID yet.

Table 1: Summary statistics

year	N	female (%)	age	yrs sch	phone (%)
2004	7,825	74.8	68.3	12.1	24.2
2006	7,371	75.3	69.6	12.1	41.4
2008	7,030	74.9	70.4	12.2	41.7
2010	8,644	71.9	67.2	12.4	34.9
2012	8,232	71.6	68.2	12.5	41.5
2014	7,758	71.6	69.1	12.5	56.3
2016	8,711	70.2	67.1	12.1	63.3

The first column presesnts number of repondents. The second shows the proportion of female respondents. The third shows the avearge age, and the fourth shows the average years of schooling. The last colum presents the proportion of phone interview mode compared to face-to-face interview.

Table 2: Summary statistics for the selected questions

Question	N	MEAN	SD	DK	RF	<10
Home Val	32.901	222385.6	688993.5	16.1	1.1	.2
Prop tax	35,800	1974.7	4174.1	18.9	1.3	4.9
SSI Inc	38,748	987.6	494.8	5.9	8.9	1.9
Checking	41,551	25526.1	77946.2	11.4	10.7	3.5
Vehicle	40,658	13785.4	68174.2	18.2	1.4	2
Food home	54,625	89.5	1359.6	10.2	1	1.8
Food out	54,625	24.6	181.4	3.2	.7	40.6
OOP Doc	29,871	564.2	2441.8	18.4	.5	5.8
OOP Dent	26,997	1104.4	2075.7	7.7	.4	3.3
OOP Drug	36,339	82.7	357.4	13	.5	12.1

The last three columns are the shares. DK and RF columns present the percent of answering “Don’t Know” and “Refused to Answer”, respectively. The last column <10 presents the percent of numerical answers less than 10.

Table 3: Regression of Proxy on controls in HRS

	(1) proxy1	(2) proxy2	(3) proxy3	(4) proxy4	(5) proxy5	(6) proxy6	(7) proxy7
age	0.000477	-0.00141	-0.00402***	-0.0102***	-0.0103***	-0.00690***	-0.0136***
	(0.000856)	(0.000859)	(0.000881)	(0.000848)	(0.000818)	(0.000886)	(0.000750)
cog	0.0276***	0.0289***	0.0614***	0.0910***	0.105***	0.131***	0.101***
	(0.00652)	(0.00655)	(0.00671)	(0.00646)	(0.00623)	(0.00674)	(0.00571)
phone	-0.0591***	-0.0657***	-0.0557***	-0.0477***	-0.0208**	-0.0265**	-0.000589
	(0.0109)	(0.0110)	(0.0112)	(0.0108)	(0.0104)	(0.0113)	(0.00956)
edu	-0.0238***	-0.0301***	-0.0276***	-0.0319***	-0.0184***	0.00767***	-0.0136***
	(0.00214)	(0.00215)	(0.00220)	(0.00212)	(0.00205)	(0.00221)	(0.00188)
female	0.0957***	0.103***	0.0796***	0.0443***	-0.0328***	0.109***	-0.0530***
	(0.0123)	(0.0123)	(0.0126)	(0.0122)	(0.0117)	(0.0127)	(0.0108)
wealth Q	-0.00382***	-0.00642***	-0.00274***	-0.00279***	0.00285***	0.0115***	0.00321***
	(0.000201)	(0.000202)	(0.000207)	(0.000199)	(0.000192)	(0.000208)	(0.000176)
N	24770	24770	24873	24873	24873	24868	24873
$Ind\ FX$	No	No	No	No	No	No	No
R^2	0.033	0.072	0.024	0.036	0.037	0.178	0.054
adj. R^2	0.033	0.072	0.024	0.036	0.036	0.178	0.054
F	140.0	320.4	103.3	156.6	157.5	895.7	235.4

Standard errors in parentheses

* $p < 0.10$, ** $p < .05$, *** $p < .01$

The cognitive score and all proxies are standardized.

The detail information on proxies are presented in Figure 6.

Table 4: Regression of Cognitive score on Proxies

	(1)	(2)	(3)
	cog	cog	cog
age	-0.0361***	-0.0360***	-0.0339***
	(0.000802)	(0.000802)	(0.000804)
phone	0.0280***	0.0282***	0.0275***
	(0.0106)	(0.0106)	(0.0106)
edu	0.0997***	0.0999***	0.0975***
	(0.00199)	(0.00200)	(0.00199)
female	0.288***	0.288***	0.281***
	(0.0118)	(0.0118)	(0.0118)
wealth Q	0.00643***	0.00650***	0.00502***
	(0.000193)	(0.000195)	(0.000214)
proxy 1	0.0263***		
	(0.00619)		
proxy 2		0.0272***	-0.00301
		(0.00616)	(0.00649)
proxy 6			0.0855***
			(0.00697)
proxy 7			0.0750***
			(0.00822)
N	24770	24770	24756
$Ind\ FX$	No	No	No
R^2	0.248	0.249	0.261
adj. R^2	0.248	0.248	0.261
F	1364.6	1364.9	1092.8

Standard errors in parentheses

* $p < 0.10$, ** $p < .05$, *** $p < .01$

All proxies are standardized.

The detail information on proxies are presented in Figure 6.

Table 5: Regression of Proxy on controls in HRS

	(1) proxy1	(2) proxy2	(3) proxy3	(4) proxy4	(5) proxy5	(6) proxy6	(7) proxy7
age	-0.00309	-0.00771***	-0.00723***	-0.0150***	-0.00905***	-0.0167***	-0.0140***
	(0.00219)	(0.00216)	(0.00227)	(0.00214)	(0.00213)	(0.00232)	(0.00191)
cog	0.0231*	0.0250**	0.0400***	0.0526***	0.0519***	0.0435***	0.0498***
	(0.0125)	(0.0122)	(0.0127)	(0.0118)	(0.0119)	(0.0121)	(0.0106)
phone	-0.0427***	-0.0453***	-0.0370***	-0.0300**	-0.0190	-0.0290**	0.000573
	(0.0138)	(0.0135)	(0.0140)	(0.0127)	(0.0127)	(0.0131)	(0.0113)
wealth Q	-0.00102**	-0.00252***	0.000570	0.00162***	0.00458***	0.00712***	0.00534***
	(0.000516)	(0.000512)	(0.000514)	(0.000503)	(0.000520)	(0.000529)	(0.000494)
N	24770	24770	24873	24873	24873	24868	24873
$Ind\ FX$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.535	0.578	0.539	0.585	0.558	0.658	0.590
adj. R^2	0.290	0.355	0.296	0.367	0.326	0.479	0.375
F	5.514	14.29	10.77	30.15	35.54	79.97	55.71

Standard errors in parentheses

* $p < 0.10$, ** $p < .05$, *** $p < .01$

Standard errors are clustered at the individual level for the fixed effects model.

The cognitive score and all proxies are standardized.

The detail information on proxies are presented in Figure 6.

Table 6: Robustness Check I using Proxy 4

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Pro tax	SSI	checking	food at home	OOP doc	OOP dent	drug
age	-0.0157***	-0.0123***	-0.0169***	-0.0138***	-0.0143***	-0.0157***	-0.0114***
	(0.00222)	(0.00231)	(0.00225)	(0.00208)	(0.00213)	(0.00218)	(0.00210)
cog	0.0530***	0.0387***	0.0557***	0.0543***	0.0456***	0.0489***	0.0513***
	(0.0123)	(0.0128)	(0.0122)	(0.0115)	(0.0118)	(0.0120)	(0.0115)
phone	-0.0276**	-0.0189	-0.0290**	-0.0285**	-0.0348***	-0.0293**	-0.0309**
	(0.0132)	(0.0140)	(0.0133)	(0.0125)	(0.0127)	(0.0129)	(0.0126)
wealth Q	0.00163***	0.000868	0.00276***	0.00126**	0.00129**	0.00153***	0.00145***
	(0.000533)	(0.000543)	(0.000516)	(0.000489)	(0.000503)	(0.000514)	(0.000496)
N	24864	24845	24869	24797	24873	24870	24871
$Ind\ FX$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.557	0.532	0.585	0.592	0.581	0.576	0.591
adj. R^2	0.324	0.286	0.367	0.377	0.361	0.354	0.376
F	29.29	14.94	38.08	27.62	27.15	29.76	23.10

Standard errors in parentheses

* $p < 0.10$, ** $p < .05$, *** $p < .01$

Standard errors are clustered at the individual level for the fixed effects model.

The cognitive score and all proxies are standardized.

The detail information on proxies are presented in Figure 6.

Table 7: Robustness Check II with home and vehicle values

	(1) proxy1	(2) proxy2	(3) proxy3	(4) proxy4	(5) proxy5	(6) proxy6	(7) proxy7
age	-0.00283	-0.00570***	-0.00605***	-0.0116***	-0.00966***	-0.0217***	-0.0115***
	(0.00202)	(0.00195)	(0.00214)	(0.00198)	(0.00197)	(0.00214)	(0.00176)
cog	0.0210*	0.0212*	0.0345***	0.0445***	0.0509***	0.0541***	0.0451***
	(0.0116)	(0.0111)	(0.0120)	(0.0109)	(0.0111)	(0.0111)	(0.00977)
phone	-0.0463***	-0.0504***	-0.0380***	-0.0279**	-0.0129	-0.0220*	0.0101
	(0.0126)	(0.0120)	(0.0130)	(0.0115)	(0.0116)	(0.0118)	(0.0102)
wealth Q	-0.00102**	-0.00277***	0.00102**	0.00217***	0.00635***	0.00874***	0.00637***
	(0.000483)	(0.000473)	(0.000487)	(0.000465)	(0.000487)	(0.000485)	(0.000450)
N	24853	24853	24879	24879	24879	24878	24879
$Ind\ FX$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.548	0.607	0.556	0.612	0.588	0.712	0.614
adj. R^2	0.311	0.401	0.324	0.409	0.372	0.561	0.411
F	6.675	16.97	10.72	26.95	62.45	144.0	74.52

Standard errors in parentheses

* $p < 0.10$, ** $p < .05$, *** $p < .01$

Standard errors are clustered at the individual level for the fixed effects model.

The cognitive score and all proxies are standardized.

The detail information on proxies are presented in Figure 6.

Table 8: Robustness Check III with Age 50+

	(1) proxy1	(2) proxy2	(3) proxy3	(4) proxy4	(5) proxy5	(6) proxy6	(7) proxy7
age	0.00791***	0.00244	0.00677***	-0.00131	0.00413***	0.00620***	-0.00562***
	(0.00158)	(0.00153)	(0.00165)	(0.00152)	(0.00151)	(0.00172)	(0.00131)
cog	0.0216**	0.0232**	0.0371***	0.0468***	0.0474***	0.0507***	0.0417***
	(0.00947)	(0.00922)	(0.00975)	(0.00886)	(0.00891)	(0.00912)	(0.00777)
phone	-0.0368***	-0.0387***	-0.0321***	-0.0248***	-0.0161*	-0.0305***	0.00237
	(0.0105)	(0.0101)	(0.0108)	(0.00955)	(0.00970)	(0.0100)	(0.00820)
wealth Q	-0.000854**	-0.00248***	0.000416	0.000924**	0.00401***	0.00634***	0.00432***
	(0.000397)	(0.000389)	(0.000405)	(0.000380)	(0.000391)	(0.000411)	(0.000353)
N	41157	41157	41332	41332	41332	41354	41332
$Ind\ FX$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.532	0.575	0.527	0.563	0.524	0.645	0.563
adj. R^2	0.308	0.372	0.302	0.354	0.297	0.475	0.355
F	10.17	15.06	8.670	11.27	34.94	71.31	50.54

Standard errors in parentheses

* $p < 0.10$, ** $p < .05$, *** $p < .01$

Standard errors are clustered at the individual level for the fixed effects model.

The cognitive score and all proxies are standardized.

The detail information on proxies are presented in Figure 6.

Table 9: Robustness Check IV with demotivation

	(1) proxy1	(2) proxy2	(3) proxy3	(4) proxy4	(5) proxy5	(6) proxy6	(7) proxy7
age	-0.00335	-0.00803***	-0.00874***	-0.0177***	-0.0122***	-0.0193***	-0.0176***
	(0.00219)	(0.00217)	(0.00224)	(0.00205)	(0.00202)	(0.00223)	(0.00175)
cog	0.0222*	0.0239*	0.0335***	0.0408***	0.0385***	0.0322***	0.0341***
	(0.0125)	(0.0122)	(0.0126)	(0.0113)	(0.0112)	(0.0115)	(0.00941)
phone	-0.0428***	-0.0454***	-0.0378***	-0.0314**	-0.0206*	-0.0303**	-0.00135
	(0.0138)	(0.0135)	(0.0139)	(0.0123)	(0.0120)	(0.0127)	(0.0103)
wealth Q	-0.00124**	-0.00278***	-0.000830	-0.000945*	0.00167***	0.00467***	0.00192***
	(0.000529)	(0.000527)	(0.000521)	(0.000488)	(0.000491)	(0.000515)	(0.000451)
motiv	-0.719**	-0.859**	-4.518***	-8.261***	-9.412***	-7.917***	-11.04***
	(0.366)	(0.359)	(0.350)	(0.355)	(0.367)	(0.327)	(0.368)
N	24770	24770	24873	24873	24873	24868	24873
$Ind\ FX$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.535	0.578	0.547	0.614	0.599	0.679	0.655
adj. R^2	0.291	0.356	0.309	0.411	0.388	0.510	0.475
F	5.146	12.30	43.53	133.8	156.3	189.2	213.2

Standard errors in parentheses

* $p < 0.10$, ** $p < .05$, *** $p < .01$

Standard errors are clustered at the individual level for the fixed effects model.

The cognitive score and all proxies are standardized.

The detail information on proxies are presented in Figure 6.

Table 10: Regression of Proxy on controls in PSID

	(1) proxy1	(2) proxy2	(3) proxy3	(4) proxy4	(5) proxy5	(6) proxy6	(7) proxy7
age	0.00880***	0.00698**	0.00831***	0.00517*	0.00183	0.00436*	-0.00475***
	(0.00275)	(0.00279)	(0.00282)	(0.00277)	(0.00242)	(0.00252)	(0.00176)
female	0.0489	0.0865*	0.0354	0.0562	-0.106**	0.0390	-0.0607*
	(0.0506)	(0.0514)	(0.0518)	(0.0511)	(0.0446)	(0.0464)	(0.0324)
edu	-0.0401***	-0.0416***	-0.0376***	-0.0342***	-0.0194***	0.0118*	0.0108**
	(0.00697)	(0.00707)	(0.00714)	(0.00703)	(0.00613)	(0.00639)	(0.00445)
married	-0.176***	-0.195***	-0.182***	-0.193***	-0.146***	0.257***	-0.0183
	(0.0478)	(0.0485)	(0.0490)	(0.0482)	(0.0421)	(0.0439)	(0.0306)
wealth Q	-0.00378***	-0.00562***	-0.00349***	-0.00475***	0.00150***	0.0103***	0.00146***
	(0.000662)	(0.000672)	(0.000678)	(0.000667)	(0.000582)	(0.000607)	(0.000423)
N	2894	2894	2897	2897	2897	2901	2897
$Ind\ FX$	No	No	No	No	No	No	No
R^2	0.080	0.115	0.069	0.087	0.008	0.182	0.020
adj. R^2	0.079	0.113	0.068	0.085	0.006	0.181	0.019
F	50.49	74.73	43.03	54.93	4.763	128.9	12.07

Standard errors in parentheses

* $p < 0.10$, ** $p < .05$, *** $p < .01$

All proxies are standardized.

The detail information on proxies are presented in Figure 6.

Table 11: Regression of Proxy on controls in PSID

	(1) proxy1	(2) proxy2	(3) proxy3	(4) proxy4	(5) proxy5	(6) proxy6	(7) proxy7
age	-0.00682	-0.00609	-0.00546	-0.00441	-0.00406	0.000200	-0.0000517
	(0.00785)	(0.00787)	(0.00798)	(0.00783)	(0.00775)	(0.00759)	(0.00621)
married	-0.0447	-0.0784	-0.0247	-0.0444	0.0991	0.575***	0.0401
	(0.197)	(0.205)	(0.193)	(0.193)	(0.207)	(0.184)	(0.192)
wealth Q	0.00164	0.000469	0.00222	0.00127	0.00398*	0.00568**	0.00169
	(0.00195)	(0.00200)	(0.00194)	(0.00194)	(0.00206)	(0.00284)	(0.00154)
N	2916	2916	2919	2919	2919	2923	2919
$Ind\ FX$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.698	0.720	0.691	0.703	0.575	0.714	0.600
adj. R^2	0.491	0.528	0.481	0.501	0.284	0.519	0.328
F	0.492	0.248	0.564	0.235	1.332	4.375	0.404

Standard errors in parentheses

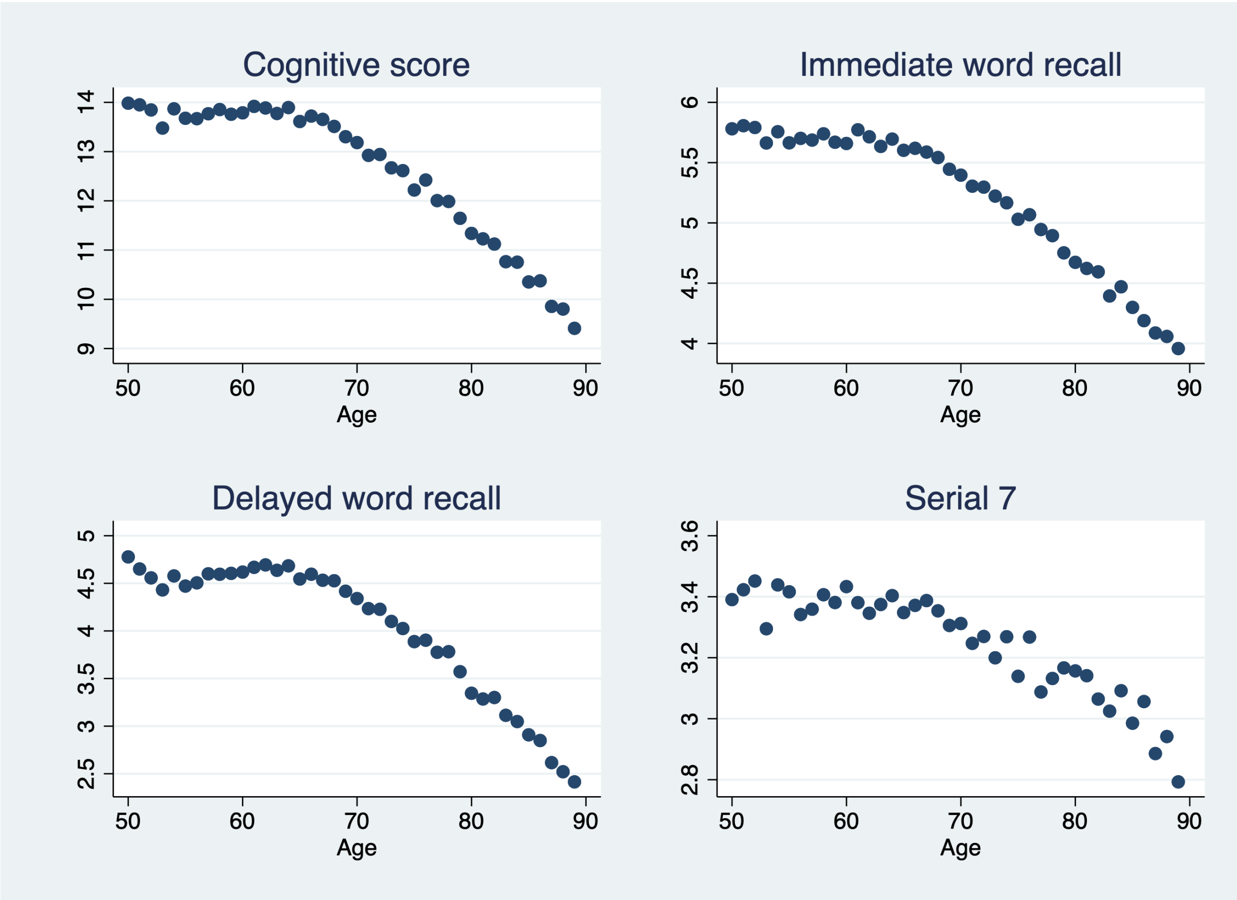
* $p < 0.10$, ** $p < .05$, *** $p < .01$

Standard errors are clustered at the individual level for the fixed effects model.

All proxies are standardized.

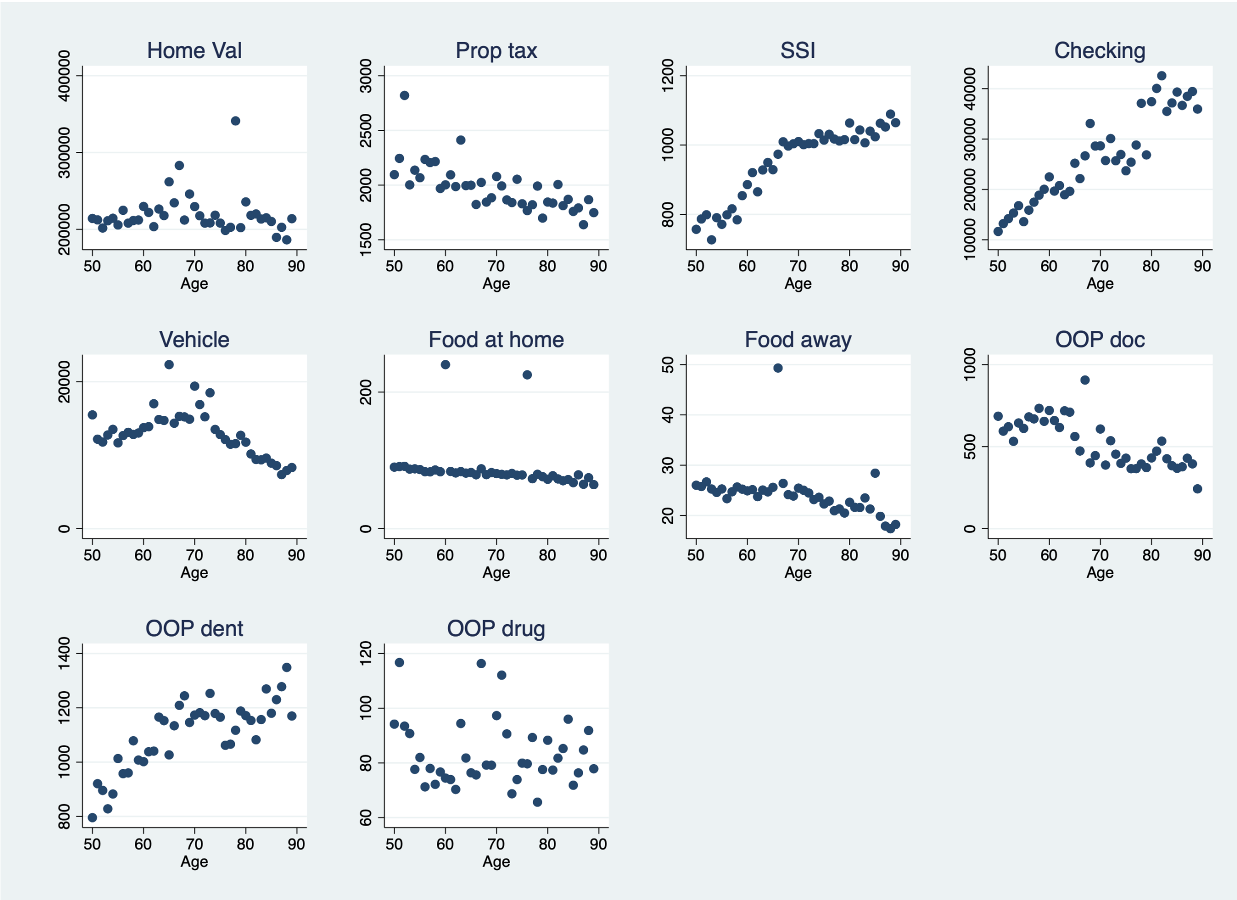
The detail information on proxies are presented in Figure 6.

Figure 1: Age trend of Cognitive score



The first panel represents the distribution of cognitive score among those who have less than 12 years of schooling. The second panel represents those who have high school degree. The thrid panel represents those who have some years in college. The last panel represents those who have more than college education.

Figure 2: Local average of numerical response



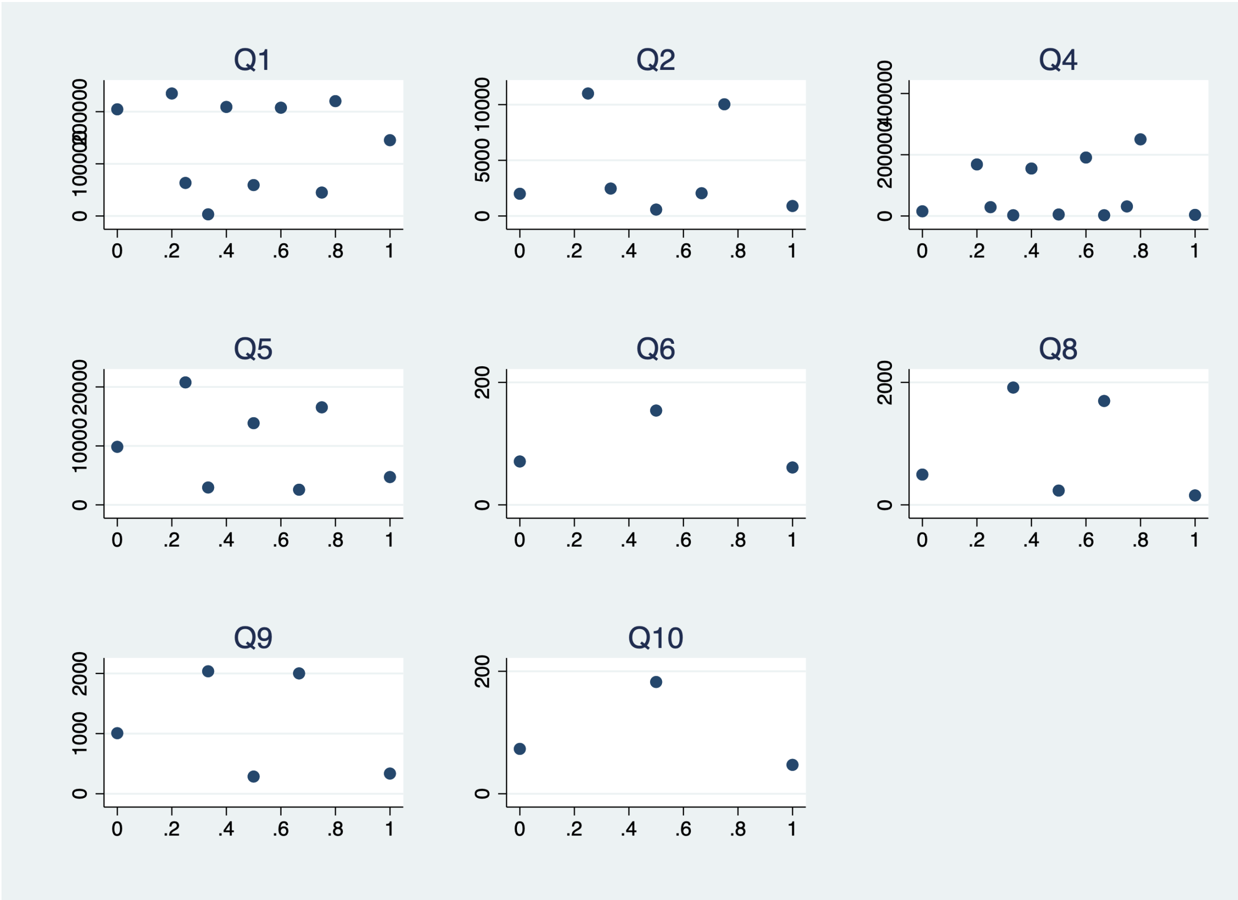
Each panel presents local average of numerical response by age. Opt-outs are ignored.

Figure 3: Response heaping

Figure 4: Distribution of First Index

In X-axis, 0 implies the maximal rounding, and 1 implies there is no trailing zero in the response.

Figure 5: Local Average of numerical response by First Index



In X-axis, 0 implies the maximal rounding, and 1 implies the most detailed numerical response. For this figure, we control for the outlier by removing 99 percentile responses.

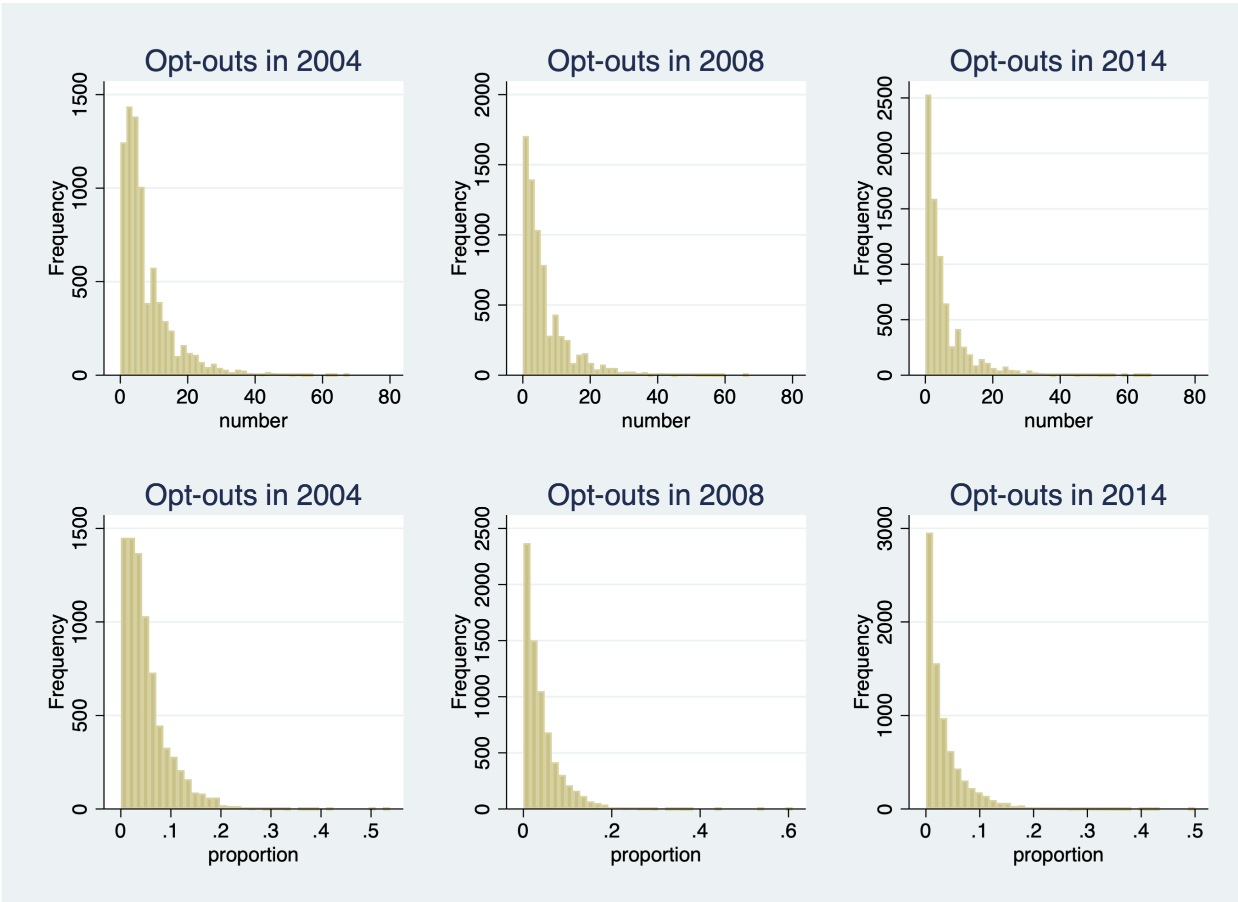
Figure 6: Local Average of Proxies by Age

proxy 1: Gideon method with reverse order $\left(\frac{n-1}{m-1}\right)$
proxy 2: $\frac{n}{m}$
proxy 3: Based on proxy 1, assign 0 to DK
proxy 4: Based on proxy 2, assign 0 to DK
proxy 5: assign 0 to DK; 1 to maximal rounding; 2 to other numerical response
proxy 6: assign 0 to DK or maximal rounding; 1 to other numerical response
proxy 7: assign 0 to DK; 1 to other numerical response
n: the number of non-trailing zeros
m: the number of total digits
The cognitive score and all proxies are standardized.

Figure 7: Local Average of Proxies by Age controlling for Individual Effect

The cognitive score and all proxies are standardized. The detail information on proxies are presented in Figure 6.

Figure 8: Distribution of Opt-outs



The first row shows the number of optouts in 2004, 2008 and 2014. The second row shows its proportion of optouts to all other responses.

Figure 9: PSID: Local Average of Proxies by Age

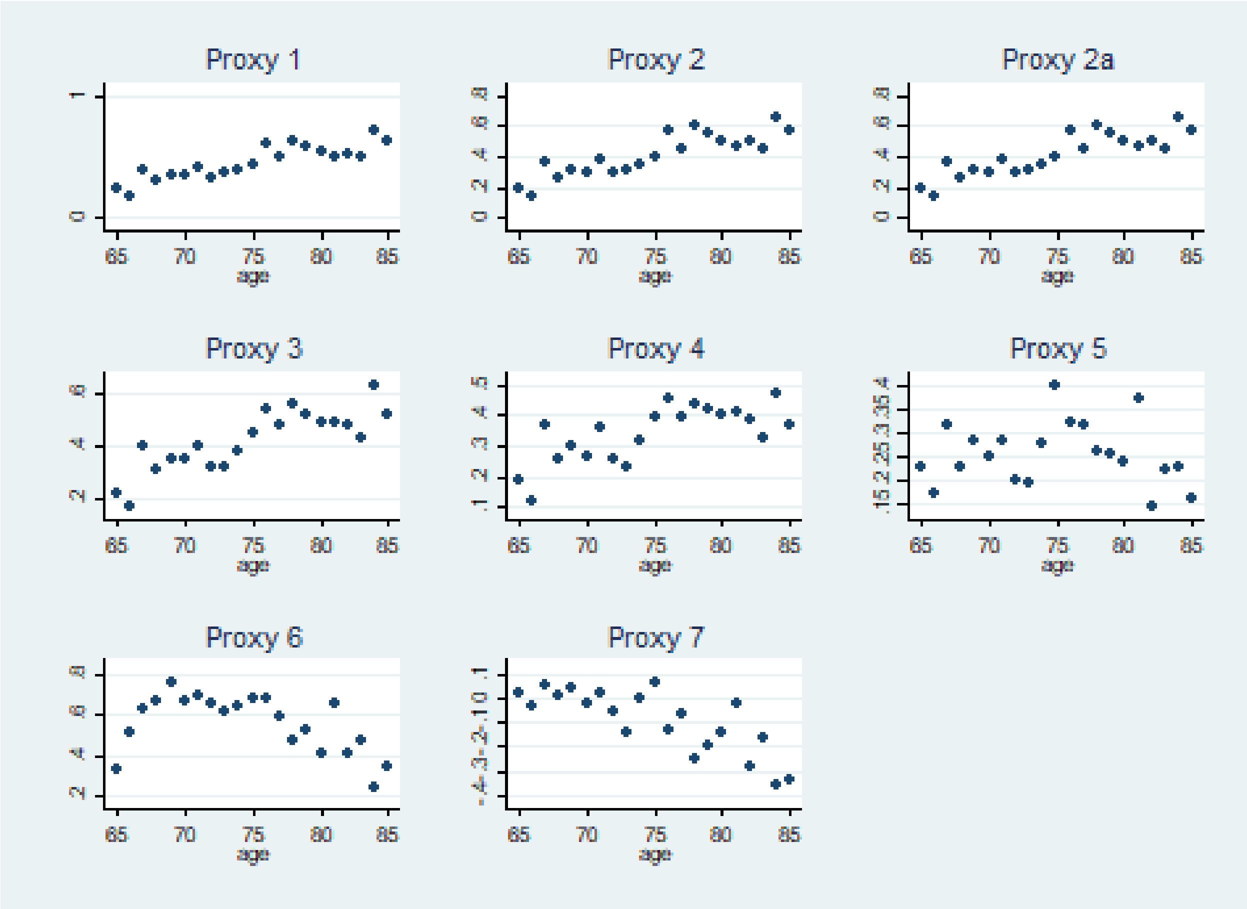
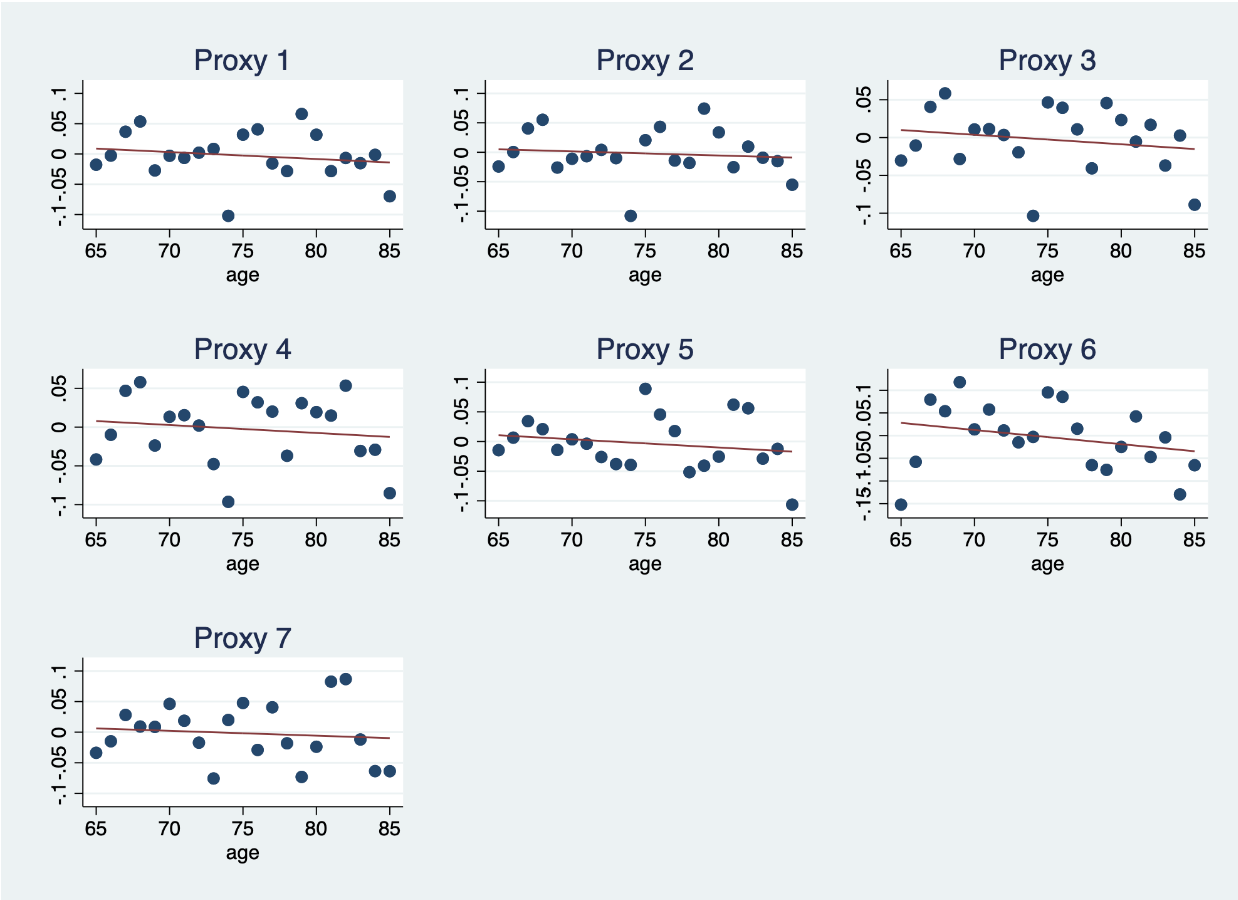


Figure 10: Local Average of Proxies by Age controlling for Individual Effect



References

Almond, Douglas. 2006. “Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population.” *Journal of political Economy* 114 (4): 672–712.

Andrews, Frank M, and A Regula Herzog. 1986. “The quality of survey data as related to age of respondent.” *Journal of the American Statistical Association* 81 (394): 403–410.

Battistin, Erich, Raffaele Miniaci, and Guglielmo Weber. 2003. “What do we learn from recall consumption data?” *Journal of Human Resources* 38 (2): 354–385.

Boyle, Phelim P, and Cormac Ó Gráda. 1986. “Fertility trends, excess mortality, and the Great Irish Famine.” *Demography,* 543–562.

Budd, John W, and Timothy Guinnane. 1991. “Intentional age-misreporting, age-heaping, and the 1908 Old Age Pensions Act in Ireland.” *Population Studies* 45 (3): 497–518.

Colsher, Patricia L, and Robert B Wallace. 1989. “Data quality and age: health and psychobehavioral correlates of item nonresponse and inconsistent responses.” *Journal of Gerontology* 44 (2): P45–P52.

Facts, Alzheimer’s Disease, and Figures 2018. 2018. “Alzheimer’s Association.” *Alzheimer’s & Dementia* 14 (3): 367–429.

Gideon, Michael, Brooke Helppie-McFall, and Joanne W Hsu. 2017. “Heaping at round numbers on financial questions: The role of satisficing.” In *Survey research methods,* 11:189. 2. NIH Public Access.

Herzog, A Regula, and Robert B Wallace. 1997. “Measures of cognitive functioning in the AHEAD Study.” *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 52 (Special_Issue): 37–48.

Holbrook, Allyson L, Sowmya Anand, Timothy P Johnson, Young Ik Cho, Sharon Shavitt, Noel Chávez, and Saul Weiner. 2014. “Response heaping in interviewer-administered surveys: Is it really a form of satisficing?” *Public Opinion Quarterly* 78 (3): 591–633.

Jessen, Frank, Rebecca E Amariglio, Martin Van Boxtel, Monique Breteler, Mathieu Ceccaldi, Gaël Chételat, Bruno Dubois, Carole Dufouil, Kathryn A Ellis, Wiesje M Van Der Flier, et al. 2014. “A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer’s disease.” *Alzheimer’s & dementia* 10 (6): 844–852.

Karr, Justin E, Raquel B Graham, Scott M Hofer, and Graciela Muniz-Terrera. 2018. “When does cognitive decline begin? A systematic review of change point studies on accelerated decline in cognitive and neurological outcomes preceding mild cognitive impairment, dementia, and death.” *Psychology and aging* 33 (2): 195.

Knäuper, Barbel, Robert F Belli, Daniel H Hill, and A Regula Herzog. 1997. “Question difficulty and respondents’ cognitive ability: The effect on data quality.” *JOURNAL OF OFFICIAL STATISTICS-STOCKHOLM-* 13:181–199.

Krosnick, Jon A. 1991. “Response strategies for coping with the cognitive demands of attitude measures in surveys.” *Applied cognitive psychology* 5 (3): 213–236.

Manski, Charles F, and Francesca Molinari. 2010. “Rounding probabilistic expectations in surveys.” *Journal of Business & Economic Statistics* 28 (2): 219–231.

Myers, Robert J. 1976. “An instance of reverse heaping of ages.” *Demography* 13 (4): 577–580.

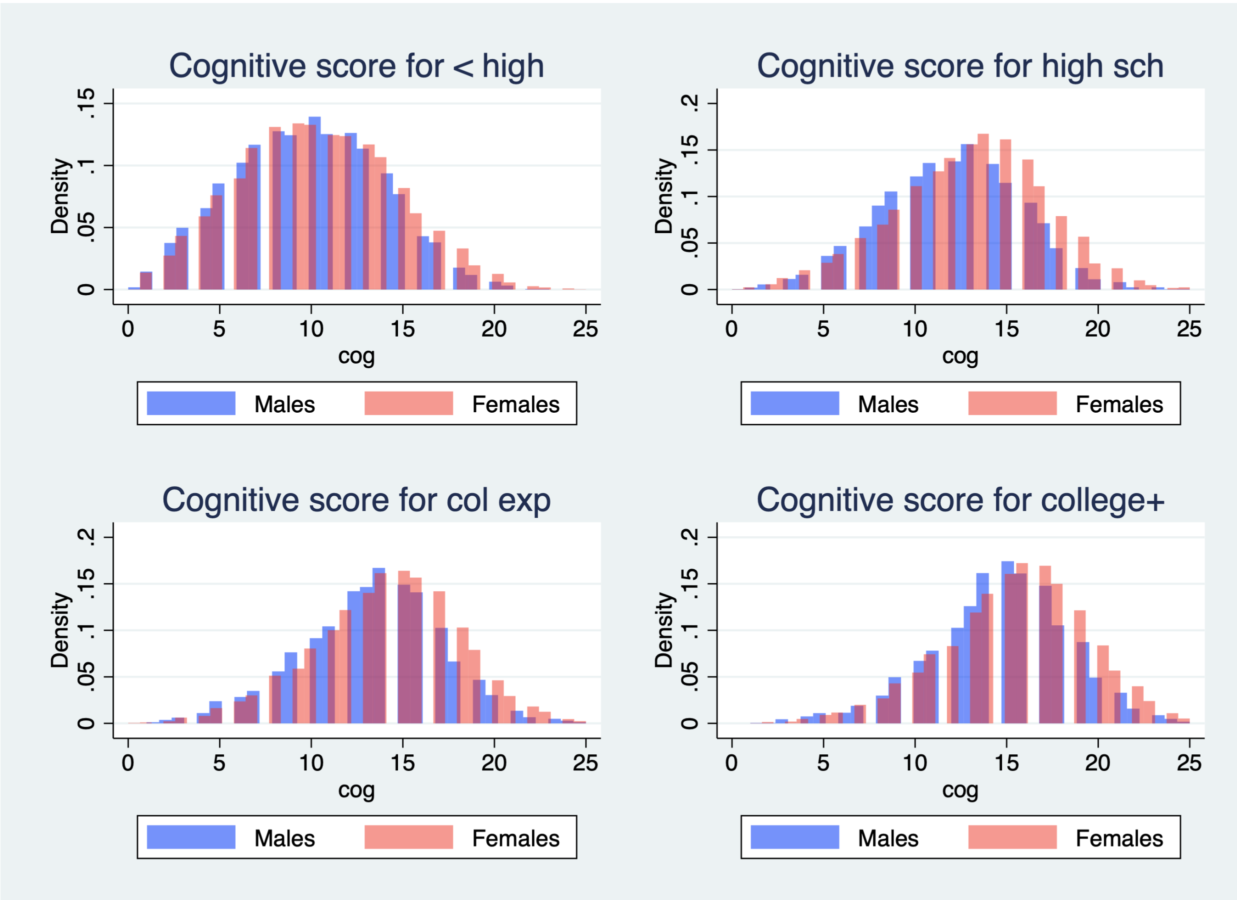
Riddles, Minsun K, Sharon L Lohr, J Michael Brick, Patrick T Langetieg, John M Payne, and Alan H Plumley. 2017. “Handling respondent rounding of wages using the IRS and CPS matched dataset.” *2016 IRS Research Bulletin,* 60–68.

Roberts, John M, and Devon D Brewer. 2001. “Measures and tests of heaping in discrete quantitative distributions.” *Journal of Applied Statistics* 28 (7): 887–896.

Zinn, S, and A Würbach. 2016. “A statistical approach to address the problem of heaping in self-reported income data.” *Journal of Applied Statistics* 43 (4): 682–703.

Appendix

Figure 11: Distribution of Cognitive Score



All panels present that female distribution is rather left skewed, and the mode of the distribution moves to the right as the level of education increases. That is, female respondents tend to have higher cognitive scores, and highly educated respondents tend to have higher cognitive scores.

Figure 12: Distribution of Second Index

In X-axis, 1 implies there is no trailing zero in the response.

Table 12: Regression of Proxy on controls in HRS with Male

	(1) proxy1	(2) proxy2	(3) proxy3	(4) proxy4	(5) proxy5	(6) proxy6	(7) proxy7
age	-0.00497	-0.0112**	-0.00954**	-0.0173***	-0.00712	-0.0150***	-0.0130***
	(0.00465)	(0.00455)	(0.00477)	(0.00437)	(0.00434)	(0.00448)	(0.00381)
cog	0.0611**	0.0625**	0.0760***	0.0808***	0.0735***	0.0762***	0.0486**
	(0.0253)	(0.0249)	(0.0262)	(0.0239)	(0.0244)	(0.0250)	(0.0213)
phone	-0.0776***	-0.0774***	-0.0642**	-0.0427*	-0.0278	-0.0507*	0.0170
	(0.0268)	(0.0264)	(0.0273)	(0.0251)	(0.0259)	(0.0262)	(0.0223)
wealth Q	-0.000900	-0.00251**	0.000269	0.00137	0.00431***	0.00787***	0.00530***
	(0.00109)	(0.00107)	(0.00112)	(0.00110)	(0.00116)	(0.00109)	(0.00107)
N	6731	6731	6761	6761	6761	6763	6761
$Ind\ FX$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.544	0.591	0.548	0.597	0.566	0.683	0.600
adj. R^2	0.279	0.353	0.286	0.364	0.315	0.500	0.368
F	5.158	8.484	6.959	11.37	8.013	23.88	11.43

Standard errors in parentheses

* $p < 0.10$, ** $p < .05$, *** $p < .01$

Standard errors are clustered at the individual level for the fixed effects model.

The cognitive score and all proxies are standardized.

The detail information on proxies are presented in Figure 6.

Table 13: Regression of Proxy on controls in HRS with Female

	(1) proxy1	(2) proxy2	(3) proxy3	(4) proxy4	(5) proxy5	(6) proxy6	(7) proxy7
age	-0.00248	-0.00660***	-0.00648**	-0.0142***	-0.00968***	-0.0172***	-0.0143***
	(0.00248)	(0.00245)	(0.00258)	(0.00245)	(0.00244)	(0.00270)	(0.00221)
cog	0.0121	0.0142	0.0295**	0.0444***	0.0455***	0.0337**	0.0502***
	(0.0143)	(0.0140)	(0.0145)	(0.0135)	(0.0136)	(0.0138)	(0.0121)
phone	-0.0306*	-0.0341**	-0.0277*	-0.0256*	-0.0160	-0.0212	-0.00509
	(0.0160)	(0.0157)	(0.0163)	(0.0147)	(0.0145)	(0.0151)	(0.0130)
wealth Q	-0.00108*	-0.00254***	0.000646	0.00168***	0.00466***	0.00687***	0.00535***
	(0.000586)	(0.000581)	(0.000578)	(0.000565)	(0.000579)	(0.000605)	(0.000556)
N	18039	18039	18112	18112	18112	18105	18112
$Ind\ FX$	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.530	0.570	0.533	0.579	0.555	0.648	0.587
adj. R^2	0.291	0.351	0.297	0.366	0.330	0.470	0.377
F	2.362	8.204	5.564	20.02	28.36	56.75	45.09

Standard errors in parentheses

* $p < 0.10$, ** $p < .05$, *** $p < .01$

Standard errors are clustered at the individual level for the fixed effects model.

The cognitive score and all proxies are standardized.

The detail information on proxies are presented in Figure 6.